

УДК 003

*К. Токумару
Оита, Япония*

НУЛЕВАЯ ПОГРЕШНОСТЬ ПРИ ВВОДЕ ЛИНГВИСТИЧЕСКОЙ ИНФОРМАЦИИ И КОДИРОВАНИЕ В ТЕКСТАХ ДОГЭНА (1200-1253)

Аннотация. Мы должны изучить подлинность и истинность лингвистической информации, прежде чем поместить ее в логическую схему мозга. В противном случае мы видим хаотичные результаты, а сама схема мозга загрязнена и повреждена ложной и шумной информацией. Авторы могут защитить свои оригинальные тексты посмертно, кодируя от этих различных шумов. Доген добавил к своему тексту постскрипты, чтобы будущие читатели могли идентифицировать его автограф.

Ключевые слова: интеллект человека, аутентичный текст, фальсификация, цифровая Лингвистика, теории информации, форвардная коррекция ошибок (FEC), постскрипты.

*К. Tokumaru
Oita, Japan*

ZERO ERROR REQUIREMENT FOR INCOMING LINGUISTIC INFORMATION AND POSTSCRIPT CODING BY DOGEN (1200-1253)

Abstract. We must examine the authenticity and truth of linguistic information before putting it into an in-brain logical circuit. Otherwise we see chaotic results, and the brain circuit itself is contaminated and damaged by false and noisy information. Authors can protect their original texts posthumously by coding against these various noises. Dogen added postscripts to his text so that future readers could identify his autograph.

Key words: human intelligence, authentic text, falsification, digital linguistics, information theories, Forward Error Correction (FEC), postscripts.

1 Digital Evolutions of Linguistic Humans Digital Linguistics (DL) is an interdisciplinary study that identifies human language as a digital evolution of mammal analog vocal sign communications. Analog signs are unique with their sound waveforms but limited in number, whilst human digital word signs are infinite by permutation of their logical property, phonemes. Digital system consists of autonomous logical networks to regulate information and generate complexity: computer networks and reproduction/evolution of living organism are digital. Inheritance and development of human collective/individual intelligence should also be understood as digital.

A Bricolage & 3 Breakthroughs	Logical Properties and Circuit Logic	Results & Achievement
Vertebrate Spinal Sign Reflex Mechanism	Dichotomy (non-linear) Dualism (A+Bi=C)	Sign Reflex & Learning, Thought
Laryngeal Descent for Vowel (72-66KA)	Phonemes and Morae (P&M) in Speech Sound	Infinite word signs, Grammar
Character Set (5KA)(phonogram / ideogram)	Externalized, Shared and Long-term P&M Memory	Abstract concept, Civilization
Computer Networks (Now)	Electronically Networked Interactive P&M Database	Keyword search & Ubiquitous access

Table-1 A Bricolage and 3 Breakthroughs for Digital Evolution of Linguistic Humans

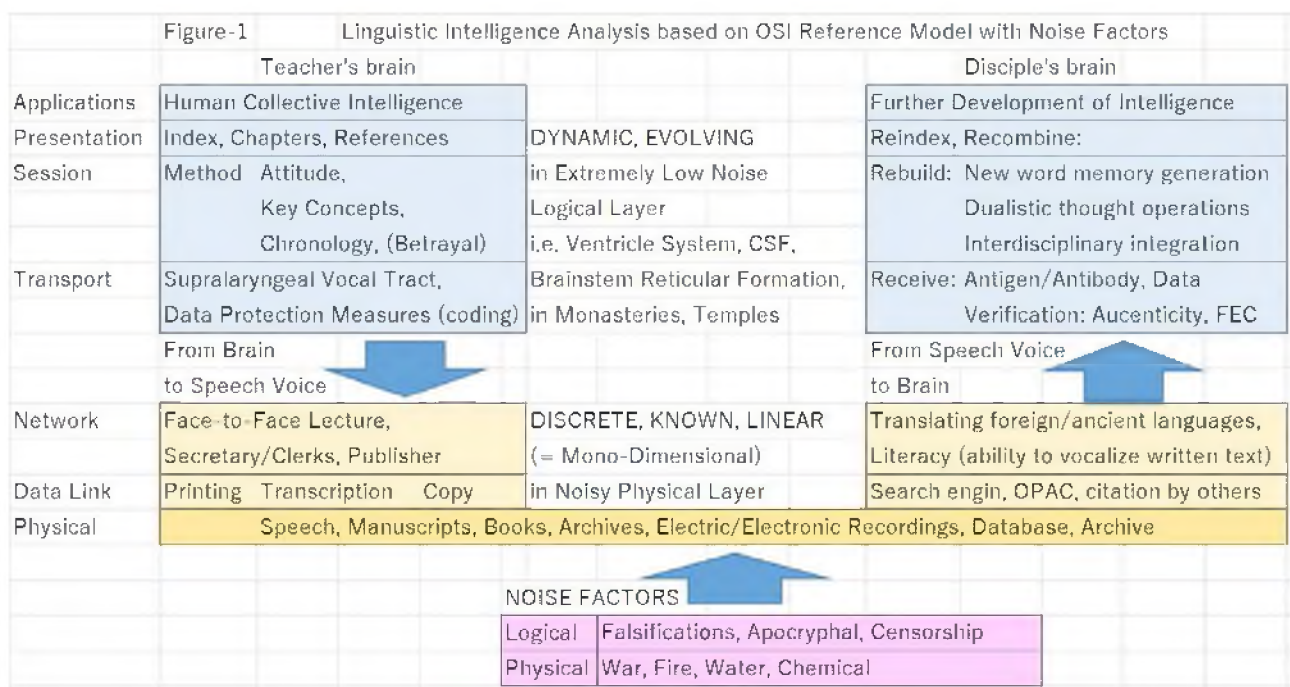
DL hypothesized that word signs are processed by vertebrate spinal sign reflex mechanism, which is neuro-immune cell idiotypic (antigen-antibody) networks inside the ventricle system. [Tokumaru 2017a] DL also identified that there are three unique evolutions for linguistic humans: laryngeal descent for vowel vocalization (66 KA in South Africa), invention of a character set (5 KA in Mesopotamia) and computer networks (Now). Laryngeal descent was an anatomical evolution, but character set and computer networks exist outside of our body, and one has to be trained and skilled to make use of them. The complexity of intelligence increases synergistically at each level.

The sign reflex mechanism is an unconscious self-protection and life-support mechanism, and some of its characteristics are not suitable for linguistic processing and intelligence. We should be aware of its molecular/cellular level mechanisms and networking phenomena to master its use and to overcome its restrictions for further development of the individual/collective intelligence.

2 The Character Set Gave Birth To Civilization It should be recognized that the character set was invented not in Africa nor in Europe, but in Mesopotamia, a very vast flat land, where earth and sand sediment filled the sea, between the Eurasia and Gondwana continents. The size of land was beyond the perception of humans at that time and some form of recording system was needed.

It can be concluded that the character set was not autopoietically invented like grammar. And, it was not the brain of linguistic humans which needed to have an external memory system. It was invented to correspond to the administrative requirements to govern unimaginably vast areas, and only those who had gone through special training of orthography could read and write. Probably they had not foreseen that a character set should help develop a civilization.

A civilization is linguistic phenomena. At the end of their biological life, linguistic persons write their accumulated knowledge with a character set, so that subsequent generations can share their thoughts and experiences. This linguistic phenomenon enabled rapid and serial innovations which we call Civilization.



3 Error Zero Requirement in Logical Layer A schematic diagram is helpful when we think of complex system. Figure-1 is the author's original combination model of Open System Interconnection Reference Model for computer networks and General Communication Model, which displays linguistic information transfer from a teacher to a pupil taking noise factors into account. [TanenbaumWeatherall2008] [Shannon 1948]

According to J.v. Neumann [1951], the guiding principle to understand digital system is "signal-to-noise ratio" in classical communication theory. Figure-1 displays the contrast of very low noise environment in logical layer which surprizingly enables dynamic phenomena, and the noise factors in the physical layer to be overcome.

J. Piaget [1947] said, "Logic is the mirror of thought, and not vice versa." The dynamic level of human intelligence depends on the in-brain logical circuit for conceptual sign processing. And as we are born innocent or "tabula rasa", we need to construct this logic for linguistic processing postnatally as efficient and as precise as possible. It is plausible that logic is based on an individual concept device and its networking memories, to be accumulated through learning and thought operations. Thus, "Any error can vitiate the result in its entirety. Thus they are permitted not a single error. ... No error should occur anywhere in the entire procedure" for individual linguistic processing, and, inter alia, logical circuit construction. [von Neumann 1951]

In order to avoid any errors, it is necessary (i) to identify the most advanced and correct information source (i.e. a contemporary teacher or a great historical person), and (ii) to examine the authenticity of every input data from that person. In information theories, errors are divided into two categories, source coding errors (SCE) for which an author is responsible, and channel coding errors (CCE) for which he is not responsible. As these two consist of an excluded middle, when both SCE and CCE are corrected, we can get error-free information. If one can identify the most advanced, correct and accessible teacher and follow his authentic words, those in later generations can easily go beyond the goal of his teacher, and thus civilization and sciences thrive.

4 Source Coding Errors Authors are responsible for source coding errors such as errors in premises, experimental methods, observation and arguments, etc. However such errors can be overcome if the original author honestly and faithfully recorded the fact. With linguistic information, we can reproduce the experiment and reconsider it in reference to today's scientific knowledge, which brings us breakthroughs for new interpretations and hypotheses.

Authors are responsible to provide all the necessary information to verify what they claim, and they must be honest and faithful to the fact. These are the requirements for not only scientific papers but also literature. In this context, authors committing plagiarism and using ghostwriters should be regarded as betrayers against truth and their texts should be disregarded, as they are unable to provide any reliable or useful data for productive re-examination.

Readers must carefully read word by word and between the lines to confirm the author's honesty in his literary style, consistent and careful wording, clear and ambitious purpose, etc. To become familiar with the author's style, his other books and papers, biography or autobiography, oral history and manuscripts, etc. should also be referred to. Their contemporaries' witness and literature, i.e. autobiography of the author's

spouse, oral history of his colleagues, etc., may also provide useful information.

5 Channel Coding Errors and Forward Error Correction (FEC) CCE are generated by natural and artificial noises. Digital information in the noisy channel takes a mono-dimensional shape consisting of mutually-distinctive and known signals such as 4 RNAs of AGUC in messenger RNA, binary 0/1 bit in computer networks and phonemic syllables in linguistic communications. Physical errors are random and a function of temperature, and countermeasures can be taken as redundancy: degeneracy in codon – amino acid translation, error correction code and onomatopoeitic etymology.

Falsification and apocryphal are logical and artificial noises. As readers in the future have no way to contact authors in the afterlife, there is no retroactive way to differentiate an author's authentic signal from falsifying noises. From the readers' side it is impossible to verify if the words are representing the original author's will or not. Only printed articles, papers and books published during an author's lifetime with his own proofreading are supposed to be authentic.

In this regard, Forward Error Correction (FEC) is fascinating. FEC is an error correction technique by an information receiver without having to contact the information sender. The information sender analyzes the data and sends the analysis with data, then the information receiver implements similar analysis to determine if there are any errors occurred during transmission and, when there are any, to correct them based on the logical integrity of the analysis result. FEC is one of the most important technologies in computer networks to achieve a zero-error requirement.

Can FEC be applied to linguistic information? First, can we analyze linguistic text and extract any logical properties? Yes, we can count the numbers of characters, sentences, paragraphs, chapters, particular words/characters, poems, etc. In addition, each text bears its own logical properties such as author's name, clerk's name, date and place of lecture, date and place of clean copy, etc. And, serial numbers can be assigned to individual volumes so that addition and deletion become visible.

Second, how can such logical properties be attached to the text and sent together? Postscripts can be added after the text and is a way to convey some essential facts related to the text.

Third, how can future readers recognize some postscripts function as FEC, and differentiate between original and fake? When they read the postscript, it is necessary to see if the author needed to protect his text, and if he implied the use of FEC. Then that postscript should carefully be examined as an error correction code.

6 Contradiction among Dogen's text An example of such postscript coding was discovered in a Japanese text of the 13th century.

Dogen (1200-1253) was a Japanese scientist, philosopher and Buddhist monk, known for his extensive and voluminous work on Buddhism philosophy, *Shobogenzo*, and his analects, *Dogen Osho Koroku*. "The Complete Work of Dogen with Translation into Contemporary Japanese" published in the 21st century consisting of 17 volumes, among which 9 are for *Shobogenzo* and 4 for *Dogen Osho Koroku*.

However study on his work has not advanced partly because of its volume and partly of the contradictions among his writings. For example, there is a 75 volume version of *Shobogenzo*, and a "new edition" of 12 volumes of *Shobogenzo*. And the postscript of the 12th volume in the new edition says, "The Master had had an intention to rewrite the old version, and to make new 100 volumes version. However, because of his illness, he completed the 12th volume and passed away." Readers are bewildered by knowing that the new version is only completed 12 volumes and that Dogen himself did not like the 75 volume version. To date, scholars have been confused and have nothing to do with such contradictions.

7 Postscript of *Shobogenzo* At the end of each volume of the *Shobogenzo* 75 volume version, Dogen added, as a postscript, combined with the series title "shobogenzo", a volume title "Genjokoan", and serialized number "No.1", like "Shobogenzo Makahannyaharamita No.2", "Shobogenzo Bussho No.3", which is followed by the date and place of delivery, date and place of clean copy, etc. In some cases, Dogen assigned a special number to a particular volume. For example, "Shobogenzo Busso (lineage of the Buddha) No.52" is serialized as 52 although it was delivered at the early stage of his activities in January 1241. (Figure-2) It seems that Dogen had the intention to declare he was the 52nd descendent of Buddha.

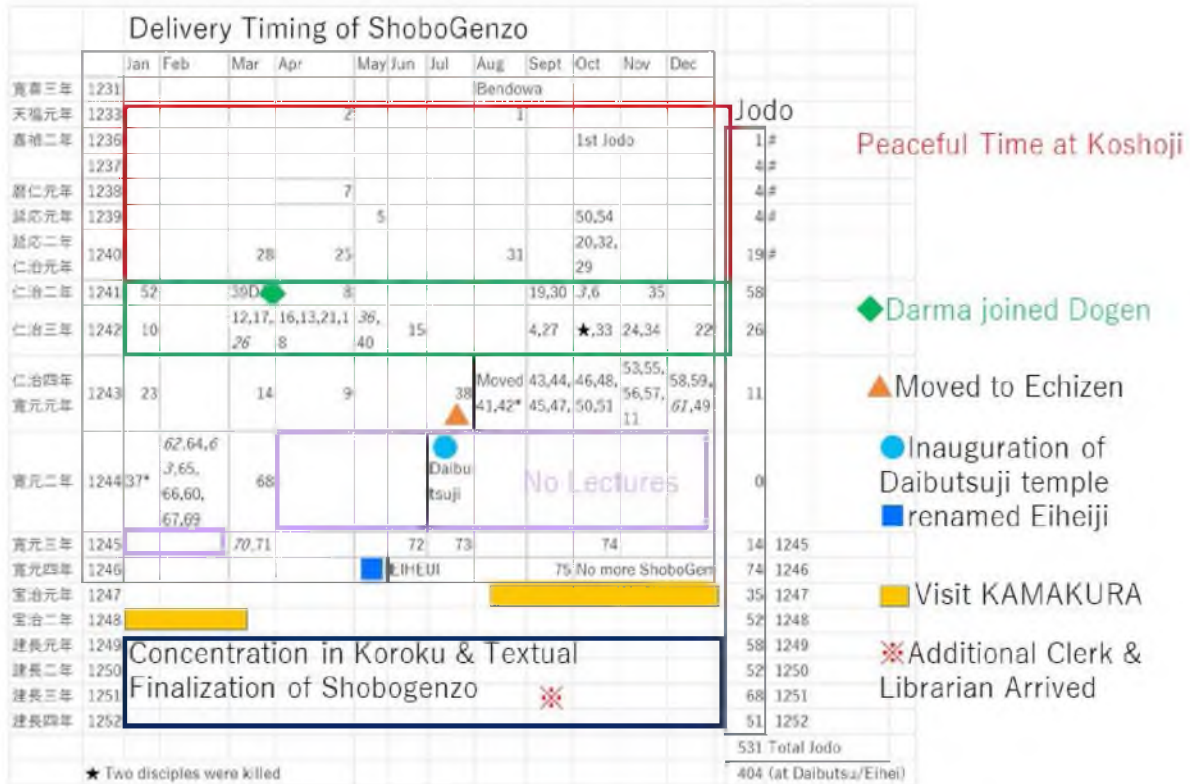
On the other hand, postscripts of the "New 12 Volume Version" are insufficient to proclaim their authenticity and contain contradictions within themselves. (Table-3) They don't indicate a delivery date, which means that they were not delivered by Dogen. The serial numbers are not assigned chronologically. Why did the author of the postscript of the new 12th version say that Dogen had finished only up to the 12th volume? Is it a lie? Most clean copy dates are after the demise of Dogen, on 28th August 1253. Apparently Dogen did not proofread them. Readers should disregard these 12 volumes as Dogen's work.

8 Relationship between Dogen and his disciples *Dogen Osho Koroku* revealed that there were severe conflicts between Dogen and his disciples from the Dharma sect, who joined Dogen's school in March 1241 and seem to have the intention to take over the Dogen's school. These Dharma disciples did not sincerely study nor practice, and Dogen scolded them in Jodo No. 134 in 1245.

Apparant differences in teaching between the Dogen and Dharma sects are shown in Jodo No. 91, Dogen said "Don't say that our sect has no word" No word is the teaching of Dharma sect. He declared instead, "Our sect is only words" in Jodo 128. He probably observed the disciples' reactions to his words, he said that

his current concern is “if one person tells a lie, all the following persons will take it as true,” (Jodo No. 131) highlighting his concern on falsification by his disciples. Dogen also said in Jodo No. 193, “Never change my teaching”, which indicates his concern with alteration of his texts by his disciples. If he needed, he should have been looking for any effective method to protect his text after his death.

Did he have a chance to protect his text by using postscripts? There is a precedence of using postscripts to point out the authenticity of the text, in “Goshui Wakashu” (Anthology of Waka Poem) in 1086 by Michitoshi Fujiwara, a member of the Fujiwara family, to which Dogen belonged.



Postscript of Shobogenzo 75 Volumes (Vol.1-15)					
S/N		Delivery	Place	Clean Copy	Place
1	GenjoKoan	1233.8.15		Included in 1252	
2	MakaHannyaHaramita	1233.4	Koshoji	1244.3.21	Kippoji
3	Bussho	1241.10.14	Koshoji	1258.4.25	
4	ShinshinGakudo	1242.9.9	Koshoji	1243.1.2	
5	SokushinZebutsu	1239.5.25	Koshoji	1245.7.12	Daibutsuji
6	Gyobutsulgi	1241 Oct	Koshoji		
7	IkkaMeiju	1238.4.18	Koshoji	1243.7.23	Kippoji
8	ShinFukatoku	1241.4	Koshoji		
9	Kobutsushin	1243.4.29	Rokuhara	1244.5.12	Kippoji
10	Taigo	1242.1.28	Koshoji	1244.3.20	Kippoji
11	Zazengi	1243 Winter	Kippoji		
12	Zazenshin	1242.3.18	Koshoji		
13	KaiinZanmai	1242.8.20	Koshoji	1243	
14	Kuuge	1243.3.10	Koshoji	1244.1.27	Kippoji
15	Komyo	1242.6.2	Koshoji	1244.12.13	Daibutsuji

Table-2 Shobogenzo 75 Vol. Ver. PS Data

Post Script of so-called 12 volume version			
	Title	Delivery Date	Clean Copy Date
No.1	Shukke Kudoku	Not Available	Aug 6, 1310
No.2	Jukai	Not Available	Not Available
No.3	Kesa Kudoku		May 25, 1275 *
No.4	Hotsu Bodaishin	Not Available	April 9, 1255
No.5	Kuyou Shobutu	Not Available	June 23, 1279
No.6	Sanpou Kie	Not Available	May 21, 1279
No.7	Shinjin Inga	Not Available	April 15, 1255
No.8	Sanjigou	Not Available	March 9, 1253
No.9	Sima	Not Available	April 15, 1255
No.10	Shizenbiku	Not Available	April 15, 1255
No.11	Ippyakuhachihoumyoumon	Not Available	Not Available
No.12	Hachidai-Ningaku	Not Available	Jan 6, 1253

Table-3 Shobogenzo New 12 Vol. Ver. PS Data

9 Indication of Error Correction Code If Dogen had invented any unusual method to protect his text, he should make it clear it so that the future reader could understand the meaning of his postscripts as an error correction code. Jodo No. 473, the third lecture in the last volume for Jodo lectures, seems to be such. "Dharma said that my laws shall remain unchanged 8,000 years later even a piece of hair in the same way during my living days. And my Buddha said that he should leave small amount of hair to protect his disciples inheriting his law after his death. He also said that he should leave twenty years of his life in the world to benefit his disciples. Today, by chance, I have a poor poem by myself as follows : Plum flowers in December receive moon light. On top of thickly laid snow in snowy mountain I added frost. Buddha's face still remains there even now. And benefit remote disciples significantly."

The author interprets "thickly laid snow" as a metaphor of his work. At the end of Jodo 471, he also used this metaphor: "In paintings you may see often, but last night it was furious snow storm." If snow is his work, "I added frost" can metaphorize "he protected text with code". In No.473, Dogen discusses Dharma and Buddha for their posthumous protection of law, "frost" can be interpreted as a protection method.

10 Conclusion: Error Zero Requirement and Information Verification DL identified that linguistic humans use sign reflex mechanism for linguistic processing. As it is reflexive, it cannot take time to investigate the incoming signs. All or none response should be OK for evacuation and mating. But for complex and sophisticated human linguistic digital communications, it is obligatory to verify the authenticity of incoming information and correct it if there are source coding or channel coding errors. While to date many people have discussed "Epokhe" in Greek philosophy in its various meaning, the author proposes to interpret it as "stop reflexive reaction, and examine words carefully". The new civilization shall start when linguistic humans pay full respect and attention to complex sophisticated language.

References

1 Piaget, J. *La Psychologie de l'intelligence*, Paris, Armand Collin 1947

2 Shannon C. E. *A Mathematical Theory of Communication*, *The Bell System Technical Journal*, Vol. 27, pp. 379-423, 623-656, 1948.

3 Tanenbaum A. S., Wetherall D. J. (2011) *Computer Networks (5 Ed) Prentice Hall*

4 Tokumaru K., *Psychology, logics and molecular/cellular biology for the mechanism of concept and grammar: inside ventricle system immune cell network hypotheses*, *Cognitive Studies of Language Vol. XXX September 20-22, 2017a*, *Russian Aca. Sci.* pp860-864, 2017a

5 Tokumaru K., *Dogen wo Yomitoku*, *Fuzambo Int'l*, 2017b von Neumann, J. *The General and Logical Theory of Automata* 1951.